

BULLIED, Neil J.  
Appl. No. 10/632,847  
June 11, 2007

**AMENDMENTS TO THE SPECIFICATION:**

Please amend the paragraph beginning at page 5, line 1, as follows:

In PCT/GB96/02122 (WO-A-97/08311) the disclosure of which is incorporated by reference we have disclosed that specific regions within the C-terminal propeptide are the recognition sequences involved in the specificity of association between C-terminal propeptide domains of pro- $\alpha$  chains during the formation of procollagens. These recognition sequences were identified as having the following amino acid sequences for each respective pro- $\alpha$  chain:

|                       |  |
|-----------------------|--|
| pro- $\alpha$ 1(I)    | GGQGSDPADV AIQLTFLRLM STE ( <u>SEQ ID NO:1</u> ) |
| pro- $\alpha$ 2 (I)   | NVEGVTSKEM ATQLAFMRLL ANY ( <u>SEQ ID NO:2</u> ) |
| pro- $\alpha$ 1 (II)  | GDDNLAPNTA NVQMTFLRLL STE ( <u>SEQ ID NO:3</u> ) |
| pro- $\alpha$ 1 (III) | GNPELPEDVLDVQLAFLRLL SSR ( <u>SEQ ID NO:4</u> )  |
| pro- $\alpha$ 1 (V)   | VDAEGNPVG.VQMTFLRLL SAS ( <u>SEQ ID NO:5</u> )   |
| pro- $\alpha$ 2 (V)   | GDHQSPNTAI .TQMTFLRLL SKE ( <u>SEQ ID NO:6</u> ) |
| pro- $\alpha$ 1 (XI)  | LDVEGNSINM .VQMTFLKLL TAS ( <u>SEQ ID NO:7</u> ) |
| pro- $\alpha$ 2 (XI)  | VDSEGSPVG.VQLTFLRLL SVS ( <u>SEQ ID NO:8</u> )   |

Please amend the paragraph beginning at page 13, line 1, as follows:

Figure 2 shows an alignment plot of the C- terminal propeptide domains of pro- $\alpha$  chains from type I (SEQ ID NO:16 and SEQ ID NO:17) and III (SEQ ID NO:18) collagen. The alignment shows amino acids which are identical (#) or those which are conserved (-). The

BULLIED, Neil J.  
Appl. No. 10/632,847  
June 11, 2007

conserved cysteine residues are numbered 1-8, while letters A, B, C, F, G denote the first amino acid at the junctions between pro $\alpha$ 1 (III) chains and pro $\alpha$ 2(I) chains of the Example;

Please amend the paragraph beginning at page 16, line 1, as follows:

Figure 13 is a schematic representation of sequence alignment of the chain selectivity recognition domains in other fibrillar procollagens, sequence homology within the 23 residue B-G motif is illustrated (SEQ ID NOs:1-8, respectively), the boxed regions indicating the position of the unique 15 residue sub-domain which directs pro- $\alpha$  chain discrimination.

Please amend the paragraph beginning at page 17, line 21, as follows:

p $\alpha$ 1(III) $\Delta$ 1 and p $\alpha$ 2(I) $\Delta$ 1 are recombinant pro- $\alpha$  chains with truncated c chain domains which have been described previously (see Lees and Bulleid (1994) J. Biol. Chem. 269 p24354-24360 1994). Chimaeric molecules were generated by PCR overlap extension using the principles outlined by Horton (1993) Methods in Molecular Biology Vol 15, Chapter 25, Humana Press Inc., Totowa, NJ. PCRs (100 $\mu$ l) comprised template DNA (500 ng), oligonucleotide primers (100 pmol each) in 10 mM KCl, 20 mM Tris-HCl pH 8.8, 10mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 2 mM MgSO<sub>4</sub>, 0.1% (v/v) Triton X-100, 300  $\mu$ M each dNTP. Ten rounds of amplification were performed in the presence of 1 unit Vent DNA polymerase (New England Biolabs, MA). Recombinants p $\alpha$ 2(I) $\Delta$ 1:(III)CP, A, F, S<sup>S-C</sup>, C<sup>S-C</sup> were generated using a 5' oligonucleotide primer (5' 'AGATGGTCGCACTGGACATC 3') (SEQ ID NO:9) complementary to a sequence 70 bp upstream of an *Sfi*I site in p $\alpha$ 2(I) $\Delta$ 1 and a 3' oligonucleotide primer (5' TCGCAGGGATCCGTCGGTCACTTGCACTGGTT 3') (SEQ ID NO:10) complementary to a region 100 bp downstream to the stop codon in p $\alpha$ 1 (III) $\Delta$ 1. A *Bam*HI site was introduced into

BULLIED, Neil J.  
Appl. No. 10/632,847  
June 11, 2007

this primer to facilitate subsequent sub-cloning steps. Pairs of internal oligonucleotides, of which one included a 20 nucleotide overlap, were designed to generate molecules with precise junctions as delineated (see Figs 2 and 3) Overlap extension yielded a product of ~990 bp which was purified, digested with *XhoI*-*Bam*HI and ligated into  $\text{p}\alpha 2(\text{I})\Delta 1$  from which a 1080 bp *XhoI*-*Bam*HI fragment had been excised. Recombinants  $\text{p}\alpha 1(\text{III})\Delta 1:(\text{I})\text{CP,C}$  were synthesized in a similar manner using a 5' oligonucleotide (5' AATGGAGCTCCTGGACCCATG 3') (SEQ ID NO:11) complementary to a sequence 100bp upstream of an *XhoI* site in a  $\text{p}\alpha(\text{III})\Delta 1$  and a 3' amplification primer (5' CTGCTAGGTACCAAATGGAAGGATTCAGCTTT 3') (SEQ ID NO:12) which incorporated a *KpnI* site and was complementary to a region 100bp downstream of the stop codon in  $\text{p}\alpha 2(\text{I})\Delta 1$ . Overlap extension produced a fragment of 1100 bp which was digested with *XhoI* and *KpnI* and ligated into  $\text{p}\alpha 1(\text{III})\Delta$  from which an 1860 bp fragment had been removed. Recombinant  $\text{p}\alpha 2(\text{I}):(\text{III})\text{BGR}$  was constructed using the same amplification primer used to synthesize the  $\text{pro}\alpha 2(\text{I})\Delta 1:(\text{III})$  series of chimeras and a 3' oligonucleotide which was identical to that used to generate the  $\text{pro}\alpha 1(\text{III})\Delta 1:(\text{I})\text{CP,C}$  constructs except that it contained a *Bam*HI site instead of *KpnI* (both complementary to  $\text{p}\alpha 2(\text{I})\Delta 1$ ). Primary amplification products were generated from  $\text{p}\alpha 2(\text{I})\Delta 1:(\text{III})\text{B}^{\text{S-c}}$  and  $\text{p}\alpha 2(\text{I})\Delta 1$  with internal oligonucleotides determining the junction. Overlap extension produced a fragment which was digested with *SfiI* and *Bam*HI and ligated into  $\text{p}\alpha 2(\text{I})\Delta 1$ . Site directed mutagenesis was performed essentially as described by Kunkel et al. (Kunkel et al. (1987) Methods in Enzymol. 154 p 367-382), except that extension reactions were performed in the presence of 1 unit T4 DNA polymerase and 1  $\mu\text{g}$  T4 gene 32 protein (Boehringer, Lewes, UK).

BULLEID, Neil J.  
 Appl. No. 10/632,847  
 June 11, 2007

Please amend the paragraph beginning at page 25, line 18, as follows:

Procollagen chain selectivity is probably mediated through one or more of the variable domains located within the C-terminal propeptide domain. The sequence between the B- and C-junctions is one of the least conserved among the procollagen C-propeptides (Figure 2), yet to inventors have demonstrated that inclusion of this domain, in the absence of  $\text{pro}\alpha 1(\text{III})$  sequence distal to the C-junction, is not sufficient to direct chain assembly. To ascertain whether the recognition sequence for chain recognition had indeed been interrupted a further recombinant,  $\text{pro}\alpha 2(\text{I}):(\text{III})\text{BGR}^{\text{S}\rightarrow\text{C}}$  (B-G replacement) was generated, which contained all of the  $\text{pro}\alpha 1(\text{I})\Delta 1$  sequence apart from the Ser  $\rightarrow$  Cys mutation at Cys2 and a stretch of 23 amino acids derived from the type III C-propeptide which spans the C-junction from points B to G, the B-G motif:  $^b\text{GNPELPEDVLDV}^{\text{C}}\text{QLAFLRLSSR}^{\text{S}}$  (SEQ ID NO:4) (underscoring indicates the most divergent residues, see Figure 2). The location of the G-boundary in the replacement motif allowed for the inclusion of the first non-conserved residues after the C-junction (SR). When expressed in the presence of SP-cells the chimeric  $\text{pro}\alpha 2(\text{I}):(\text{III})\text{BGR}^{\text{S}\rightarrow\text{C}}$  chains were able to form inter-chain disulfide-bonded molecules (Figure 8, lane 6) demonstrating that the C-terminal propeptide domains were capable of self-association. Furthermore, this hybrid was able to fold and form a stable triple-helix as judged by the formation of a protease-resistant fragment (Figure 9, lane 3).  $\text{Pro}\alpha 2(\text{I}):(\text{III})\text{BGR}^{\text{S}\rightarrow\text{C}}$  contains a Ser  $\rightarrow$  Cys substitution which enabled the inventors to assay for the formation of disulfide-bonded trimers. Previous data demonstrated that this substitution alone does not enable wild-type  $\text{pro}\alpha 2(\text{I})\Delta 1$  claims to form homotrimers (Lees and Bulleid, 1994). Nevertheless, to eliminate the possibility that this mutation influences the assembly pattern a revertant  $\text{pro}\alpha 1(\text{I}):(\text{III})\text{BGR}^{\text{C}\rightarrow\text{S}}$  which contains the wild-type complement of Cys residues was created. As expected  $\text{pro}\alpha 2(\text{I}):(\text{III})\text{BGR}^{\text{C}\rightarrow\text{S}}$  was unable to form disulfide-bonded

BULLIED, Neil J.  
Appl. No. 10/632,847  
June 11, 2007

trimers (Figure 10, lane 5) but did assemble correctly into a protease-resistant triple helix (Figure 11, lane 3). Thus, the 23-residue B-G motif contains all of the information required to direct procollagen self-assembly.

Please amend the paragraphs beginning at page 27, line 2, as follows:

Analysis of the 23 amino acid B-G motif from the  $\text{pro}\alpha 1(\text{III})$  and  $\text{pro}\alpha 2(\text{I})$  chains (Figure 13) indicates that residues 13-20 (QLAFLRLL) (SEQ ID NO:15) are identical with the exception of position 17, Leu (L) in  $\text{pro}\alpha 1(\text{III})$  and Met (M) in  $\text{pro}\alpha 2(\text{I})$ . Using site-directed mutagenesis the inventors substituted the existing Leu residue with Met to create  $\text{pro}\alpha 2(\text{I}):(\text{III})\text{BGR}^{\text{L}\rightarrow\text{M}}$  and monitored the effect of this mutation on chain assembly. The Leu  $\rightarrow$  Met mutagenesis was performed using recombinant  $\text{pro}\alpha(\text{I}):(\text{III})\text{BGR}^{\text{S}\rightarrow\text{C}}$  and  $\text{pro}\alpha 2(\text{I}):(\text{III})\text{BGR}^{\text{L}\rightarrow\text{M}}$  and were able to form interchain disulfide-bonded molecules when analysed under non-reducing conditions (Figure 10, lanes 4 and 6). Both constructs formed protease-resistant triple-helical domains (Figure 11, lanes 1 and 3). The Leu  $\rightarrow$  Met substitution did not, therefore, disrupt the process of chain selection nor did it prevent the formation of a correctly aligned triple-helix. These observations lead to the conclusion that a discontinuous sequence of 15 amino acids: (GNPELPEDVLDV.....SSR) (SEQ ID NO:13) contains all of the information necessary to allow procollagen chains to discriminate between each other and assemble in a type-specific manner.

### 3. DISCUSSION

The molecular mechanism which enables closely related procollagen chains to discriminate between each other is a central feature of the assembly pathway. The initial

BULLIED, Neil J.  
Appl. No. 10/632,847  
June 11, 2007

interaction between the C-terminal propeptide domains both ensures that the constituent chains are correctly aligned prior to nucleation of the triple-helix and propagation in a C- to N-direction, and that component chains associate in a collagen type-specific manner. As a consequence, recognition signals which determine chain selectivity are assumed to reside within the primary sequence of this domain, presumably within a region(s) of genetic diversity. By generating chimeric procollagen molecules from parental 'mini-chains'  $\text{pro}\alpha 1(\text{III})\Delta 1$  and  $\text{pro}\alpha 2(\text{I})\Delta 1$  the inventors have demonstrated that transfer of the  $\text{pro}\alpha 1(\text{III})$  C-terminal propeptide domain to the naturally heterotrimeric  $\text{pro}\alpha 2(\text{I})$  molecule was sufficient to direct formation of homotrimers. Furthermore, analysis of a series of molecules in which specific sequences were interchanged from  $\text{pro}\alpha 1(\text{III})$  and  $\text{pro}\alpha 2(\text{I})$  - C-terminal propeptide domains allowed the inventors to identify a discontinuous sequence of 15 amino acids (GNPELPEDVLDV.....SSR) (SEQ ID NO:13) within the  $\text{pro}\alpha 1(\text{III})$  C-propeptide, which, if transferred to the corresponding region within the  $\text{pro}\alpha 1(\text{III})$  recognition motif to the  $\text{pro}\alpha 2(\text{I})$  chain did not appear to have an adverse effect on chain alignment, allowing the triple-helical domains to fold into a protease-resistant conformation. This sequence motif is, therefore, both necessary and sufficient to ensure that procollagen chains discriminate between each other and assemble in a type-specific manner.

Please amend the paragraph beginning at page 28, line 9, as follows:

In order to establish a structure-function relationship for the chain recognition domain, the inventors examined the hydropathy profile and secondary structure potential of the 23-residue B-G sequence: GNPELPEDVLDVQLAFLRLSSR (SEQ ID NO:4). The data indicate that the 15-residue chain recognition motif: GNPELPEDVLDV.....SSR (SEQ ID NO:14) is markedly hydrophilic, in contrast to the hydrophobic properties of the conserved region:

BULLIED, Neil J.  
Appl. No. 10/632,847  
June 11, 2007

QLAFLRLLL (SEQ ID NO:15). These features are entirely consistent with a potential role for this motif in mediating the initial association between the component procollagen monomers. An examination of the 15-residue recognition motif from other fibrillar procollagens predicts that they are all relatively hydrophilic and probably assume a similar structural conformation, regardless of the degree of diversity in the primary sequence (Figure 13). It is, presumably, the nature of the amino acids changes which provides the distinguishing topographical features necessary to ensure differential chain association. An examination of the B-G sequence alignment (Figure 13) indicates that residues 1, 2, 12 and 21 are more tightly conserved than amino acids 3-11, 22 and 23, suggesting that the latter may form a core recognition sequence that is of critical importance in the selection process. We do not know whether the other four residues participate directly in chain discrimination but this can be tested experimentally by site-directed mutagenesis.

Substitute the Sequence Listing submitted herewith for that filed December 22, 2004.